PHAB toxins: A unique family of predatory sea anemone toxins evolving via intra-gene concerted evolution defines a new peptide fold

Bruno Madio[1], Steve Peigneur[3], Yanni K. Y. Chin[1], Brett R. Hamilton[2, 4], Sónia Troeira Henriques[1], Jennifer J. Smith[1], Ben Cristofori-Armstrong[1], Zoltan Dekan[1], Berin A. Boughton[5], Paul F. Alewood[1], Jan Tytgat[3], Glenn F. King[1*] and Eivind A.B. Undheim[2*]

[1]*Institute for Molecular Bioscience,* [2]*Centre for Advanced Imaging, and* [4]*Centre for Microscopy and Microanalysis, The University of Queensland, St Lucia, Queensland, 4072, Australia*

[3]*Toxicology and Pharmacology, University of Leuven, Leuven 3000, Belgium*

[5]*Metabolomics Australia, School of Biosciences, The University of Melbourne, Parkville, Victoria, 3010, Australia*

*Correspondence: e.undheim@uq.edu.au (lead contact) and glenn.king@imb.uq.edu.au

Keywords*:* Neurotoxin, ion channel, mass spectrometry imaging, 3D structure, concerted evolution, extreme resolution mass spectrometry imaging, on-tissue reduction alkylation

Abstract

Sea anemone venoms have long been recognized as a rich source of peptides with interesting pharmacological and structural properties, but they still contain many uncharacterized bioactive compounds. Here we report the discovery, three-dimensional structure, activity, tissue localization, and putative function of a novel sea anemone peptide toxin that constitutes a new, sixth type of voltage-gated potassium channel (K_V) toxin from sea anemones. Comprised of just 17 residues, κ-actitoxin-Ate1a (Ate1a) is the shortest sea anemone toxin reported to date, and it adopts a novel three-dimensional structure that we have named the Proline-Hinged Asymmetric β-hairpin (PHAB) fold. Mass spectrometry imaging and bioassays suggest that Ate1a serves a primarily predatory function by immobilizing prey, and we show this is achieved through inhibition of *Shaker*-type K_V channels. Ate1a is encoded as a multi-domain precursor protein that yields multiple identical mature peptides, which likely evolved by multiple domain duplication events in an actinioidean ancestor. Despite this ancient evolutionary history, the PHAB-encoding gene family exhibits remarkable sequence conservation in the mature peptide domains. We demonstrate that this conservation is likely due to intra-gene concerted evolution, which has to our knowledge not previously been reported for toxin genes. We propose that the concerted evolution of toxin domains provides a hitherto unrecognized way to circumvent the effects of the costly evolutionary arms race considered to drive toxin gene evolution by ensuring efficient secretion of ecologically important predatory toxins.

Introduction

Venoms are complex cocktails of bioactive molecules that disrupt the physiology of envenomated prey [1, 2]. Although these toxins include a wide range of molecules, such as proteins, peptides, polyamines, and salts, the impressive molecular diversity of most invertebrate venoms is due to disulfide-rich peptides [3]. In animals that rely on venom for prey capture, diet and foraging ecology are thought to be major drivers of toxin evolution, with the acquisition of resistance in prey countered by diversifying selection acting on toxin genes in the predator [4]. As a result, the venoms of predatory animals tend to

be highly diverse, often containing hundreds to thousands of unique bioactive toxins [3]. One such group is sea anemones, which are benthic, sessile cnidarians that use venom for a variety of ecological functions, including prey capture, defence, digestion, and inter- and intraspecific competition.

Given the ecological importance of venom in sea anemones, and the fact that the cnidarian venom system has been evolving for >700 million years [5], it is not surprising that sea anemones have evolved a rich variety of venom toxins including enzymes, cytolysins, and neurotoxins[6, 7]. Of these, disulfide-rich peptide neurotoxins constitute the largest molecular diversity. According to the classification system proposed by Mikov and Kozlov [8], at least 17 different peptide folds have been identified in sea anemone venoms [9], although recent proteomics studies suggest that they likely contain 30 or more [10].

In addition to being the most diverse components of sea anemone venoms, neuroactive peptides are also the most well studied. They have been used as tools for probing ion channel structure and function, and for developing novel therapies [7]. For example, ShK, a venom peptide from the sea anemone *Stichodactyla helianthus*, recently completed Phase 1 clinical trials for treatment of autoimmune disease [11]. Neurotoxins from sea anemone venoms act on a diverse range of ion channels, including acid-sensing ion channels (ASIC), transient receptor potential ion (TRP) channels, and voltage-gated sodium (Na_V) and potassium (K_V) channels. Of these, K_V toxins are the most diverse group, comprising 136 of the 320 annotated sea anemone toxins in UniProtKB. These K_V toxins are currently divided into five distinct types based on their sequence, disulfide-bridge pattern, and activity [12].

Here we describe the structure, activity, function and evolution of a new, sixth type of sea anemone K_V toxin. κ-Actitoxin-Ate1a (henceforth Ate1a), from venom of the Waratah sea anemone *Actinia tenebrosa,* is the shortest sea anemone toxin reported to date, and it adopts a novel β-hairpin-like 3D fold. In contrast with many β-hairpin peptides, Ate1a lacks antimicrobial activity and instead serves a predatory function via potent inhibition of prey K_V channels. While most families of predatory toxins evolve via bursts of extensive duplication and diversification, this is not the case for the Ate1a toxin family, whose members remain remarkably well conserved despite their ancient evolution. Our data

suggest that this extreme conservation is due to intra-gene concerted evolution, a process that has to our knowledge not been previously reported for any toxin family and which we propose is a hitherto unrecognised mechanism of maintaining efficient secretion of ecologically important toxins.

Materials and Methods

Venom Collection and Fractionation

Sea anemones were housed in aquaria at The University of Queensland. Animals were kept for no longer than two months prior to any experiments, and the average conditions for the system where the animals were kept were: 10 hours light, 14 hours dark; salinity was 35.9 ppt; pH 8.22, and 29.19 °C. Venom was obtained by electrical stimulation [13] and fractionated using reverse-phase HPLC as described previously [10].

Mass Spectrometry

Peptide masses in lyophilised HPLC fractions were analysed using MALDI-TOF MS (AB SCIEX 5800 MALDI-TOF/TOF mass spectrometer). HPLC fractions were mixed 1:1 (*v*/*v*) with α-cyano-4-hydroxy-cinnamic acid (7.5 mg/mL in 50/50 acetonitrile (ACN)/H_2O, 0.1% trifluoroacetic acid. Mass spectra were collected in reflector positive mode. 1,5-Diaminonaphthalene was used as a reductive matrix [14] to sequence intact Ate1a by ISD MS. The sample was mixed 1:1 (*v*/*v*) with 1,5-diaminonaphthalene (15 mg/mL in 50/50 ACN/H_2O, 1% formic acid, and spectra were interpreted manually.

MALDI-MSI was performed according to published protocols [15, 16] using an UltraFlex III (Bruker-Daltonik, Bremen, Germany). On-tissue reduction and alkylation of cystines was carried out on de-paraffinised tissue sections using a volatile reaction protocol as described [17], but with a 3.5 mL reaction volume in a 50 mL Corning Falcon tube (Thermo Fisher Scientific).

For ultra-high mass resolution MSI we used a SolariX XR 7T FT-ICR mass spectrometer (Bruker-Daltonik, Bremen, Germany) was used and operated in the positive ion mode. Data size was set to 1M

95 across the mass range 400–6000 m/z. MALDI source was set to a laser power = 50%, a total of 500
96 shots per scan at a frequency of 2kHz, smart walk was enabled with a width of 90 µm. The Collision
97 Cell RF Frequency was set to 1.4 MHz, Collision RF Amplitude 1100 Vpp, Transfer Optics Time of
98 Flight = 1.5 ms at a frequency of 2 MHz with RF Amplitude = 400 Vpp. The sweep excitation was set
99 to 20%. For isotopic fine structure analysis, data size was set to 4M across the mass range 200–3000
100 m/z. Data was collected and averaged across 8 scans. MALDI source was set to a laser power = 50%,
101 a total of 5000 shots per scan at frequency of 2 kHz, the laser was manually moved across sample area.
102 For isolation the quadrupole was set to 1890.00 with an isolation window of 5 m/z. The Collision Cell
103 RF Frequency was set to 2 MHz and Collision RF Amplitude set to 1200 VPP. The Transfer Optics
104 were set to a Time of flight = 1.5 ms, Frequency set to 4 MHz with RF Amplitude set to 400 Vpp.
105 Sweep excitation set to 19%. For data analysis Bruker (Bruker-Daltonik, Bremen, Germany)
106 DataAnalysis 5.0 and for image analysis Bruker FlexImaging 5.0 and Bruker SCiLS Lab 2017a were
107 used.

108 *Transcriptomics*

109 Total RNA extraction by TRIzol, cDNA library preparation, transcriptome sequencing, read trimming,
110 and assembly was performed as described previously [10]. Raw sequence reads (SRA) and Trinity-
111 assembled contigs were deposited with links to BioProject accession number PRJNA414357 in the
112 NCBI BioProject database (www.ncbi.nlm.nih.gov/bioproject/). CDSs were identified using the Galaxy
113 tool '*Get open reading frames or coding sequences*' [18]. The Ate1a sequence determined using ISD-
114 MALDI-MS was used to search the translated transcriptome using NCBI BLAST+ blastp [19].

115 *Evolution of Ate1a*

116 The Ate1a prepropeptide sequence was used to search for homologues in UniProtKB, NCBI nr and
117 EST, and a tentacle transcriptome of *S. haddoni* [10] using NCBI BLAST+ blastp. Nucleotide
118 sequences were retrieved and aligned using mafft v7.304b [20] domains were extracted using CLC
119 Main Workbench v7.6.1 and maximum likelihood phylogenies reconstructed with IQ-Tree v1.5.5 [21]

for each domain type. The evolutionary model (FLU+G4) was determined using ModelFinder [22], and support values estimated by ultrafast bootstrap using 10000 iterations [23].

Ate1a Synthesis

Ate1a (H-RCKTCSKGRCRPKPNCG-NH_2) was assembled using Fmoc chemistry (0.1 mmol scale) on a Symphony automated peptide synthesiser (Protein Technologies) on Fmoc-Rink-amide polystyrene resin. Amino acid couplings, Fmoc deprotections, and removal of side-chain protecting groups were achieved using published protocols [24]. Side-chain protecting groups were: Arg(Pbf), Asn(Trt), Cys(Trt), Lys(Boc), Ser(tBu) and Thr(tBu). The crude product (ESI-MS *m/z*: calc. (avg) 632.1 $[M+3H]^{3+}$, found 632.0) was oxidatively folded by stirring in 0.1 M NH_4HCO_3 (pH 8.1) at room temperature for 3 days to give a single major isomer that was isolated by preparative HPLC (ESI-MS *m/z*: calc. (avg) 630.8 $[M+3H]^{3+}$, found 631.0).

NMR structure determination of Ate1a

The solution structure of Ate1a was determined using 2D NMR spectroscopy. Spectra of synthetic Ate1a (1 mM in 20 mM sodium phosphate, pH 6, 5% D_2O) were acquired at 10°C on a cryoprobe-equipped Avance 600 MHz spectrometer (Bruker BioSpin). Resonance assignments were made using 2D ^{1}H-^{1}H TOCSY, 2D ^{1}H-^{1}H NOESY spectra, and natural abundance 2D ^{1}H-^{15}N HSQC and ^{1}H-^{13}C HSQC spectra. Spectra were analysed using CcpNmr Analysis v2.4.1 [25]. Resonance assignments (97.8% complete) have been deposited in BioMagResBank (accession number 30342). Dihedral-angle restraints were derived using TALOS-N [26] and the restraint range set to twice the estimated standard deviation. The NOESY spectrum was manually peak-picked, then 96.3% of the peak list was automatically assigned, and structures calculated using CYANA v3.97 [27] The final structure was calculated using 66 interproton distance restraints, 6 disulfide-bond restraints and 23 dihedral-angle restraints. 200 structures were calculated, then the 20 with highest stereochemical quality as judged by MolProbity were used to represent the solution structure of Ate1a. Atomic coordinates are available from the Protein Data Bank (accession code 6AZA).

Electrophysiological characterisation of Ate1a

The pharmacological effect of Ate1a was analysed by heterologous expression of $rK_V1.1$, $rK_V1.2$, $hK_V1.3$, $rK_V1.4$, $rK_V1.5$, $rK_V1.6$, Shaker IR, $rK_V2.1$, $hK_V3.1$, $rK_V4.2$, $hK_V7.2$, $hK_V11.1$, $rNa_V1.2$, $rNa_V1.3$, $rNa_V1.4$, $hNa_V1.5$, $mNa_V1.6$, $hNa_V1.7$, $rNa_V1.8$, rASIC1a, rASIC1b, rASIC2a, and rASIC3 in *Xenopus laevis* oocytes, with lowercase *r, m* and *h* indicating the channel is of rat, mouse or human origin, respectively. The linearized plasmids were transcribed using the T7 or SP6 mMESSAGE-mMACHINE transcription kit (Ambion, Waltham, MA, USA). The $K_V1.1$ triple mutant channel was constructed as previously describes [28]. Oocytes were injected with 50 nL of cRNA at a concentration of 0.05–1 ng/nL using a micro-injector (Drummond Scientific, Broomall, PA, USA). Injected oocytes were stored at 19 °C in an ND96 solution (in mM: 96 NaCl, 2 KCl, 1.8 $CaCl_2$, 2 $MgCl_2$ and 5 HEPES; pH 7.4), supplemented with 50 μg/mL gentamicin sulfate.

Two-electrode voltage-clamp recordings were performed at room temperature (18–22 °C) using a Geneclamp 500 or Axoclamp 900A amplifier (Molecular Devices, Sunnyvale, CA, USA) controlled by a pClamp data acquisition system (Axon Instruments, Union City, CA, USA). Whole cell currents from oocytes were recorded 1–5 days after injection, when a whole cell current could be observed. The bath solution composition was ND96 (in mM: 2 NaCl, 96 KCl, 1.8 $CaCl_2$, 2 $MgCl_2$ and 5 HEPES; pH 7.4). Voltage and current electrodes were filled with KCl (3 M). The resistance of both electrodes was kept between 0.8 and 1.0 MΩ. K_V currents were filtered at 0.5 or 2 kHz using a four-pole low-pass Bessel filter, and leak subtraction was performed using a −P/4 protocol. $K_V1.1$−$K_V1.6$ and Shaker IR currents were evoked by 500 ms depolarization to 0 mV followed by a 500 ms pulse to −50 mV, from a holding potential of −90 mV. $K_V2.1$, $K_V3.1$, $K_V4.2$ and $K_V4.3$ currents were elicited by 500 ms pulses to +20 mV from a holding potential of −90 mV. Current traces of *h*ERG channels were elicited by applying a +40 mV pulse for 2.5 s followed by a step to −120 mV for 2.5 s. Na_V currents were evoked by 100 ms depolarization pulse from a holding potential of −90 mV to −20 mV, with the exception of $Na_V1.8$ which were pulsed to 0 mV. Na_V data were digitized at 20 kHz; leak and background conductance were identified by blocking channels with tetrodotoxin and subtracted from currents. ASIC currents were acquired (digitized 2 kHz and filtered at 0.01 Hz) and elicited by a drop in pH from 7.45 to 6.5, 5.5, 4.5, and 6.3 (for ASIC1a, ASIC1b, ASIC2a, and ASIC3, respectively).

Due to the lack of a $hNa_V1.1$ clone for oocyte screening, activity on this channel was assessed via manual patch-clamp electrophysiology. HEK293 cells heterologously expressing $hNa_V1.1$ (SB Drug Discovery, Glasgow, UK) were maintained at 37 °C in a humidified 5% CO_2 incubator in minimal essential medium supplemented with 10% FBS v/v, 2 mM L-glutamine and selection antibiotics as recommended by the manufacturer. Cells were grown to 70–80% confluence and passaged every 2–4 days using Detachin (Genlantis, San Diego, CA, USA). Whole cell patch-clamp recordings were obtained using a MultiClamp 700B amplifier (Molecular Devices, Sunnyvale, CA, USA). Patch pipettes were pulled from standard wall borosilicate glass capillaries (1.5 mm x 0.86 mm, OD/ID; SDR Scientific, Sydney, AUS) using a microelectrode puller (P-97; Sutter Instrument Co., Novato, CA, USA) and had a resistance of 1.2–1.5 MW when filled with pipette solution. The pipette solution was composed of (in mM) 150 CsCl, 1 EGTA, 10 HEPES and was adjusted to pH 7.2 with CsOH. The external bath solution consisted of (in mM) 140 NaCl, 4 KCl, 1 $MgCl_2$, 2 $CaCl_2$, 10 HEPES and adjusted to pH 7.4 with NaOH. Currents were monitored for at least 5 min after establishing whole-cell configuration to allow currents to stabilize. The pulse protocol consisted of cells being held at -90 mV for 10 s, followed by a hyperpolarizing step to -120 mV for 200 ms, then a depolarizing step to -15 mV for 50 ms. Series resistance and prediction compensation between 50–75% was applied to reduce voltage errors. Recorded currents were acquired with a Digidata 1550B (Molecular Devices) converter at 50 kHz after passing through a low-pass Bessel filter of 10 kHz. A P/6 subtraction protocol provided by the Clampex (Molecular Devices) acquisition software was used to remove linear leak and residual capacitance artifacts.

Ate1a concentration–response relationships data were fitted with the Hill equation $y = 100/[1 + (IC_{50}/[\text{toxin}])h]$, where y is the amplitude of the toxin-induced effect, IC_{50} is the toxin concentration at half-maximal efficacy, [toxin] is the toxin concentration and h is the Hill coefficient. In order to investigate the current–voltage (I–V) relationship, current traces were evoked by 10 mV depolarization steps from a holding potential of −90 mV. The values of I_K were plotted as function of voltage and fitted using the Boltzmann equation $I_K/I_{max} = [1 + \exp(V_g - V)/k]^{-1}$, where I_{max} represents maximal I_K, V_g is the voltage corresponding to half-maximal current and k is the

slope factor. To assess the concentration dependence of the Ate1a induced inhibitory effects, a concentration–response curve was constructed in which the percentage of current inhibition was plotted as a function of toxin concentration. Data were fitted with the Hill equation. All data represent at least three independent experiments ($n \geq 3$) and are presented as mean ± standard error. Comparison of two sample means was made using a paired Student's *t* test ($P < 0.05$). All data were analyzed using clampfit 10.3 (Molecular Devices) and origin 7.5 software (Origin Lab., Northampton, MA, USA).

Antimicrobial activity of Ate1a

Antimicrobial screening was performed by the Community for Antimicrobial Drug Discovery (CO-ADD) (www.co-add.org). Inhibition of growth was measured against five bacteria: *Escherichia coli* (ATCC 25922), *Klebsiella pneumoniae* (ATCC 700603), *Acinetobacter baumannii* (ATCC 19606), *Pseudomonas aeruginosa* (ATCC 27853) and *Staphylococcus aureus* (MRSA; ATCC 43300), and two fungi: *Candida albicans* (ATCC 90028) and *Cryptococcus neoformans* (ATCC 208821), following protocols previously described [Page 5, Supplementary Information in 29].

Susceptibility of red blood cells and cultured cells to Ate1a

Hemolytic assays employed erythrocytes isolated from fresh human blood collected from healthy donors using protocols approved by the Human Research Ethics Committee at The University of Queensland. Hemolysis was quantified as described [30] by measuring hemoglobin release at 405 nm, with melittin and cyclic gomesin as controls. The cytotoxicity of Ate1a was determined against HeLa, MCF-7, and HFF-1 cell cultures using a resazurin colorimetric assay [31], with melittin and cyclic gomesin as positive controls.

Interactions of Ate1a with lipid bilayers

Surface plasmon resonance (SPR) was used to monitor the affinity of Ate1a for lipid membranes using a Biacore 3000 instrument (GE healthcare) and an L1 chip at 25ºC. Synthetic POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) and POPS (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoserine)

(Avanti polar lipids) were used to prepare small unilamellar vesicles (SUVs, 50 nm diameter) composed of POPC or POPC/POPS (4:1 molar ratio) dispersed in HEPES buffer (10 mM HEPES containing 150 mM NaCl, pH 7.4) and homogenized by extrusion. The L1 chip possesses a dextran coat modified with alkyl chains to allow attachment of liposomes and formation of lipid bilayers. SUVs were injected onto an L1 chip for 40 min at a flow rate of 2 μL/min; the signal reached a steady state below 10,000 response units in the four flow channels confirming coverage of the chip surface and formation of stable bilayers. Serial two-fold dilutions of Ate1a, starting from 64 μM, were injected over deposited lipid bilayers for 180 s at a flow rate of 5 μL/min (association phase); dissociation was followed for 600 s [32, 33]. An N-to-C cyclized version of gomesin (UniProtKB P82358) was included for comparison. The chip was regenerated as before [34]. All solutions were freshly prepared and filtered using a 0.22 μm filter; HEPES buffer was used as running buffer. Response units were normalized to peptide-to-lipid ratio (P/L) as previously described [32].

Toxicity of Ate1a In vivo

Toxicity of Ate1a to brine shrimps (*A. salina*) and amphipods (family Talitridae) was examined as previously described [35, 36]. For assays where the toxin was dissolved in medium, synthetic Ate1a was dissolved to a concentration of 0.5 mg/mL in filtered artificial seawater. Assays were performed in 24-well plates for shrimps and 6-well plates for amphipods (Thermo-Fisher Scientific). Paralysis and lethality were assessed by microscopic observation and responsiveness to contact with a 10 μL plastic tip. Bovine serum albumin (5 mg/mL) was dissolved in the medium to be used as control (no toxicity). For injection assays, Ate1a was diluted with a physiological solution for crustaceans (in mM: NaCl 470.4, KCl 8.0, $CaCl_2$ 18.0, $MgCl_2$1.5, $NaHCO_3$ 6.0 and glucose 5.6). The injection volume was 9.4 nL. Groups of ten amphipods (4.35–11.53 mg) were challenged with 5.3 mM of toxin and observed for mortality or paralysis up to 4 h. Controls were amphipods that were not injected or were injected with crustacean physiological solution.

Results

Discovery of Ate1a

Specimens of *A. tenebrosa* were collected off the coast of North Stradbroke Island, Queensland, Australia (27°15 S, 153°15 E), and venom obtained by electrical stimulation[10]. Fractionation of venom using reversed-phase chromatography revealed a conspicuous early-eluting peak containing an unusually low-mass component (Fig. 1a), which we confirmed to be a disulfide-rich peptide by *de novo* sequencing using in-source dissociation (ISD) matrix-assisted laser desorption/ionization mass spectrometry (MALDI MS) (Fig. 1b). The toxin, which we named Ate1a, is a 17-residue peptide with two disulfide bonds and an amidated C-terminus (RCKTCSKGRCRPKPNCG-NH_2), yielding a monoisotopic mass of 1887.93 Da. Ate1a is a novel peptide with no BLAST hits in UniProtKB or NCBI databases.

To confirm the amino acid sequence of Ate1a and identify any venom homologues, we generated a transcriptome from tentacles actively regenerating venom, as described previously [10]. *De novo* assembly with Trinity yielded 87,485 contigs, which translated to 457,470 potential coding sequences (CDS). A BLAST search was used to identify the transcript encoding Ate1a, and this returned a single contig containing a partial CDS with multiple copies of a peptide domain encoding a sequence identical to that determined by ISD-MALDI-MS. Analysis of remapped reads revealed that this contig represents two unique transcripts whose CDS differ by two synonymous and two non-synonymous mutations in the propeptide regions of the Ate1a preproprotein (Figure S1, NCBI SRA accession SRR6282389).

Ate1a belongs to a novel peptide family that evolves by intra-gene concerted evolution

We used the complete sequence of the Ate1a precursor to search for related sequences. A BLAST search against the UniProtKB and GenBank nr databases returned no significant hits, but BLAST searches against the NCBI expressed sequence tag (EST) database returned one full-length and two partial Ate1a-like prepropeptide sequence in *Anemonia viridis*. Similarly, a BLAST search against our published tentacle transcriptome from *Stichodactyla haddoni* [10] yielded three unique contigs, including one full-

277 length prepropeptide. These species include two separate families (Actiniidae and Stichodactylidae),
278 suggesting that the Ate1a toxin-gene family arose in a common ancestor of the superfamily Actinioidea
279 [37].

280 All identified Ate1a-like prepropeptides are comprised of the same set of domains separated by dibasic
281 cleavage sites: a signal peptide, one or two cysteine-containing propeptide domains, and three cysteine-
282 free propeptide domains that each precede an Ate1a-like domain (Fig. 2a). The domain architecture is
283 also the same for all prepropeptides except for the position of the cysteine-containing propeptide, which
284 in Stichodactylidae (*S. haddoni*) is found as a single copy immediately following the signal peptide, but
285 in Actiniidae (*A. tenebrosa* and *A. viridis*) exists as two copies that each follow the first two Ate1a-like
286 domains. This suggests that the Ate1a precursor gene underwent early domain duplication events
287 followed by either multiple domain deletions or convergent deletions and duplications. Strikingly,
288 however, the domains share almost 100% nucleotide identity with corresponding domains within the
289 same prepropeptide (Fig. 2b). Phylogenetic analysis revealed that domains from each species form well-
290 supported clades with respect to those from other species (Fig. 2c). This suggests that the extreme
291 sequence conservation observed in domains of the Ate1a gene family is due to concerted evolution, an
292 evolutionary phenomenon that, to our knowledge, has previously only been demonstrated for a single
293 animal toxin gene family [38], but never for within-gene toxin domains.

294 *Ate1a defines a new peptide fold*

295 The unique primary structure of Ate1a prompted us to characterise its solution structure using NMR
296 spectroscopy. We synthesised Ate1a using solid phase peptide synthesis and confirmed correct folding
297 of the synthetic product by HPLC co-elution with native peptide (Fig. S2).

298 The 3D structure of Ate1a (Fig. 3) was determined using homonuclear NMR methods, and statistics for
299 the ensemble of structures are shown in Supplementary Table 1. MolProbity analysis [39] revealed that
300 the structure has excellent stereochemical quality, with no steric clashes and ~90% of residues in the
301 most favoured Ramachandran region. The precision of the structure, however, is not very high
302 (backbone RMSD 1.11± 0.28 Å), suggesting that it is highly dynamic, particularly within the longer

303 loop 3. This may be due the presence of two proline residues, which leads Ate1a to adopt two distinct
304 conformations (Fig. 3a). Both conformations adopt a fold similar to that of β-hairpin-like peptides[40],
305 where the C- and N- termini are connected via two semi-parallel disulfide bonds (C1-C4 and C2-C3).
306 One face of the toxin has a high proportion of positively charged residues (Fig. 3b) whereas the opposite
307 face is rich in hydrophobic residues (Fig. 3c).

308 Although Ate1a displays a hairpin-like structure (Fig. 4a), it is neither a true hairpin scaffold nor similar
309 to any other previously described hairpin-like peptide fold. The two disulfide-enclosed loops of Ate1a
310 differ substantially in length, with loop 1 containing just two residues compared to five in loop 3. In
311 combination with the two prolines in loop 3, this asymmetry prevents the formation of secondary
312 structures characteristic of other disulfide-enclosed hairpin-like structures such as β-hairpin
313 antimicrobial peptides (AMPs) [40] (Fig. 4b) or the cystine-stabilised α/α (CSαα) fold [41] (Fig. 4c).
314 Ate1a also differs from other hairpin-like folds found in animal toxins, such as the boundless β-hairpin
315 (BBH) [42] (Fig. 4d) and disulfide-directed hairpin (DDH) fold[43] (Fig. 4e), or the two-disulfide fold
316 of RhTx from venom of the centipede *Scolopendra subspinipes*[44] (Fig. 4f). Thus, Ate1a is the
317 prototypic member of a previously undescribed peptide fold that we coined the ***P****roline-****h****inged*
318 ***a****symmetric* ***β****-hairpin-like* (PHAB) fold.

319 *Ate1a represents a new type of sea anemone K_V-toxin*

320 Many Arg/Lys-rich, disulfide-stabilised β-hairpin peptides (e.g., gomesin and tachyplesin-1) function
321 as AMPs in the innate immune system. They often have high affinity for lipid membranes and possess
322 both anticancer and antimicrobial activity [29]. Although Ate1a does not adopt a typical β-hairpin fold,
323 it is highly positively-charged. However, Ate1a had no antimicrobial activity at concentrations up to
324 256 μg/mL. Similarly, Ate1a was not cytotoxic or cytolytic against cultured human cancer cell lines or
325 erythrocytes (Fig. S3). Consistent with these results, Ate1a displayed only weak affinity for, and rapid
326 dissociation from, lipid membranes compared to gomesin [45] (Fig. S4). Taken together, Ate1a's lack
327 of antimicrobial and cytolytic activity, as well as its low affinity for lipid membranes, suggests that it
328 does not play a role in defence against pathogens.

Ion channels are the most common molecular target of disulfide-rich venom peptides, and we therefore used electrophysiology to screen Ate1a against eight Na_V channels, twelve K_V channels, and four ASIC subtypes. Ate1a was found to selectively target several members of the *Shaker* subfamily of K_V channels; at 3 μM it inhibited currents mediated by $K_V1.1$ (84% ± 4%), $K_V1.2$ (94% ± 3%), $K_V1.3$ (38% ± 4%), $K_V1.6$ (92% ± 2%), and Shaker IR (23% ± 2%) channels (Fig. 5a). No activity was observed on other channels at the same concentration (Fig. S5). Fitting of the Hill equation to concentration–response curves for $K_V1.1$, $K_V1.2$, $K_V1.3$ and $K_V1.6$ yielded IC_{50} values of 353 nM, 146 nM, 3051 nM and 191 nM, respectively (Fig. 5b). Thus, given its unique sequence and structure, Ate1a represents a new, sixth type of sea anemone K_V toxin.

Ate1a is a toxin with a predatory function

Although the pharmacological activity of a toxin can provide clues to its ecological function, it is not by itself definitive. However, the near-universal distribution of nematocytes in the epithelium of sea anemones means that toxin function can be inferred from tissue distribution [46-48]. We therefore investigated the tissue distribution of Ate1a using MALDI-MS imaging (MSI), which allows visualization of the spatial distribution of unlabelled low mass biomolecules (1–20 kDa)[15, 16, 49]. A peak corresponding to the average mass of Ate1a was clearly observed in MALDI-TOF-MSI spectra acquired from cross-sectioned *A. tenebrosa*. The identity of this peak was further supported by on-tissue gas-phase reduction and alkylation, which resulted in a peak shift matching the alkylation of four cystines (Fig. 6a). Finally, ultra-high mass resolution analysis by MALDI-FT-ICR-MSI allowed us to fit the predicted isotope structure of Ate1a to the observed spectra and confirm its identity (Fig. 6b). MSI revealed that Ate1a is non-uniformly distributed within the body of *A. tenebrosa*, with almost exclusive localization in tentacles (Fig. 6c), suggesting that it is involved in prey capture. Ate1a mass signals were weak or absent in actinopharynx, mesenterial filaments, and gastrovascular cavity, indicating that Ate1a does not play a role in prey digestion. Moreover, *A. tenebrosa* normally retracts its tentacles in response to disturbances, and thus the weak Ate1a signal in the trunk region indicates it is not primarily involved in defence.

In addition to nematocytes, sea anemones also produce toxins in ectodermal gland cells [50]. Unlike nematocytes, which are stinging cells that inject venom, gland-cell toxins are released into the water and absorbed by prey. To determine which cell type produces Ate1a, we conducted toxicity bioassays using brine shrimp and amphipods, the latter being a major prey item of *Actinia* spp. [51, 52]. Injection of Ate1a into amphipods resulted in impaired swimming followed by contractile paralysis (Supplementary material S1). In contrast, Ate1a did not affect either species when dissolved into the medium (artificial sea water) (Supplemental videos S1 and S2). Taken together, our data suggest that Ate1a is a neurotoxin produced in nematocytes and used primarily for prey capture.

Discussion

Although sea anemone venoms are a rich source of bioactive peptides, recent omics studies have highlighted how little we still know about their composition, function, and evolution [10, 48]. Here we described the discovery and functional characterization of a new peptide class from venom of the sea anemone *A. tenebrosa*, one of the most commonly encountered sea anemones in intertidal zones around Australia and New Zealand [53]. Ate1a has a primary structure unlike any previously described peptide, and assumes a unique 3D fold that is reminiscent of β-hairpin AMPs [40].

In contrast to β-hairpin AMPs, the 3D structure of Ate1a is devoid of regular secondary structure. Instead, the asymmetry of the two sides of the β-hairpin-like structure of Ate1a prevents β-sheet formation, and distinguishes the 3D structure from previously described two-disulfide peptide folds (Fig. 4). The longer of the two 'loops' is also highly dynamic (Fig. 3a), a property facilitated by the presence of two prolines that are conserved in all identified Ate1a homologues. Proline-containing peptides have the ability to populate two discrete conformations, and this *cis-trans* conformational switch works like a hinge that can potentially serve as a precise regulator of biological function [54, 55]. While proline hinges play a diversity of roles in protein biology, one of these roles is reorienting surface loops to modulate protein binding surfaces and in turn ligand recognition [56]. Thus, we predict that the proline-hinged loop of Ate1a represents a region that is important for the function of this toxin family. Given the structural and likely functional importance of this structural feature, we named this

new structural scaffold the "*proline-hinged asymmetric β-hairpin-like*" (PHAB) fold.

Reflecting the structural distinctiveness of the PHAB fold from β-hairpin-like peptides, Ate1a does not have antimicrobial, antifungal, or cytolytic activity, nor strong affinity for lipid membranes (Fig. S2). Instead, it is a potent inhibitor of *Shaker*-type K_V channels, with nanomolar potency on $K_V1.1$, $K_V1.2$ and $K_V1.6$. K_V channels play crucial roles in neuronal signalling, muscle contraction, and secretion [57], and hence they are a common target of animal toxins. Many venomous taxa have convergently evolved toxins that target K_V channels to induce paralysis, general hyperexcitability, cardiac disorders, convulsions and death [1]. This is also the case in sea anemones, where K_V toxins are represented by five unique peptide folds: ShK (type 1), Kunitz-domain (type 2), β-defensin-like (type 3), boundless β-hairpin (type 4), and an unknown fold predicted to form an inhibitor cystine knot (type 5)[12]. The PHAB fold is unlike any of these structural scaffolds, and therefore it represents a new, sixth type of sea anemone K_V toxin.

Although Ate1a is a novel K_V toxin, correlating toxin activity with ecological function is often not straightforward [58]. However, like other cnidarians, sea anemones lack a centralised venom delivery system, and instead rely on localised production of toxins to complement their functional anatomy [46-48]. In *A. tenebrosa*, toxins are produced in five tissues and regions that have distinct ecological functions: acrorhagi (aggressive intraspecific encounters), tentacles (prey capture and immobilisation), mesenteric filaments (used principally in digestion), column (external defence after retracting tentacles), and actinopharynx (prey immobilisation and digestion). In addition, sea anemones produce toxins in two distinct cell types that deliver venom by either injection (nematocytes) or absorption following secretion into the water column (gland cells)[50]. Ate1a is found predominantly in the tentacles of *A. tenebrosa* (Fig. 6), which is suggestive of a predatory function. Moreover, Ate1a impaired swimming and led to paralysis and death when injected in amphipods, a major prey of *Actinia* species, but had no effect when dissolved into the medium. We conclude that Ate1a is a predatory toxin that cannot reach its K_V targets without being inoculated into prey by nematocysts.

Venom proteins are thought to evolve via toxin recruitment events, whereby a gene encoding a normal

407 body protein is duplicated and expressed in the venom-producing tissue [4]. Functionally important
408 toxin types are reinforced through duplication and diversification, and this is considered a hallmark of
409 toxin evolution in predatory venoms, where toxins evolve continuously to counter acquisition of prey
410 resistance [4]. Although recent research suggests that venoms evolve via a two-step process in which
411 initial rapid toxin diversification is followed by periods of purifying selection due to the metabolic costs
412 of diversifying selection [59, 60], predatory toxins nevertheless tend to be part of large, highly diverse
413 gene families. Strikingly, however, this diversity is entirely absent in the PHAB gene family, despite
414 their likely role in predation. Instead, its members are highly conserved and consist of just 2–3 almost
415 identical copies in each species (Fig. 2b).

416 The sequence conservation at the nucleotide level is not limited to between-gene copies of each species,
417 but extends to the domains encoded by each transcript. Despite the emergence and domain duplication
418 of the PHAB fold in an actinioidean ancestor, all four domain types (signal peptide, two propeptide
419 domains, and PHAB domain; Fig. 2a) are remarkably well conserved. Furthermore, the nucleotide
420 sequences encoding each domain type are more similar to the respective domains contained on the same
421 transcript than to corresponding domain copies in other species (Fig. 2c). This form of domain
422 conservation is likely to have occurred by concerted evolution, an evolutionary process driven by
423 continuous recombination that results in homogenisation of genetic variance across gene copies and so-
424 called 'horizontal evolution' [61]. Although concerted evolution has been described for a number of
425 gene families, including Na_V Type I toxins from *Nematostella vectensis* and *Actinia equina* [38], it is
426 considered rare for intra-gene protein domain repeats [62], and has never previously been reported
427 within toxin gene domains.

428 In contrast with the general view of gene duplication as a facilitator of toxin gene diversification, recent
429 studies have suggested that gene duplication may be of immediate importance for increased expression
430 levels rather than generation of sequence diversity [63]. Similarly, the concerted evolution of Na_V Type
431 I toxins from *N. vectensis* and *A. equina* has been suggested to confer a selective advantage through a
432 'dosage' effect of gene expression [38]. In the PHAB gene family, this lack of high gene-copy numbers
433 is compensated for by encoding multiple, identical toxin precursors, thereby effectively multiplying

toxin expression levels. Concerted evolution may also facilitate 'transmission' of advantageous mutations from a single toxin gene locus to other loci, or preventing the loss of highly effective toxins. Reflecting this, concerted evolution of protein domain repeats has been proposed to be triggered by arms-race type co-evolution [62].

Interestingly, similarly conserved domain repeats were also identified by Honma and colleagues [64] for the BBH-like AmeI (GenBank accession AB180685) from the venom of the sea anemone *Antheopsis maculata*, which is encoded on the same transcript as six repeats that share near-identical nucleotide sequences. Although further work is required to determine whether these domains evolve by concerted evolution, their identical nature suggests that intra-gene concerted evolution may in anemones not be restricted to PHAB toxins. Thus, concerted evolution of toxin-domain repeats may provide a hitherto unrecognised mechanism of circumventing the effects of the metabolically expensive evolutionary arms race typically considered to drive toxin gene evolution. In the case of Ate1a and other members of the PHAB family, this has led to efficient secretion through high domain copy numbers of structurally unusual but ecologically important, predatory K_V toxins.

Acknowledgments

This work was supported by the Brazilian Government (Science Without Borders PhD scholarship to BM), Australian Research Council (DECRA Fellowship DE160101142 to EABU, Future Fellowship FT150100398 to STH, ARC Linkage Grant LP140100832 to BRH and GFK), and National Health & Medical Research Council (Principal Research Fellowship APP1044414 to GFK). We thank Jason Cockington and Gillian Lawrence for maintenance of sea anemones, and Dr Alun Jones for help with mass spectrometry experiments. Antimicrobial screens were performed by CO-ADD, funded by the Wellcome Trust (UK) and The University of Queensland (Australia). We thank Dr Lachlan Rash for assistance with ASIC clones acquisition; Prof. John Wood for ASIC1a, ASIC2a, and ASIC3 clones; Prof. Stefan Gründer for the ASIC1b clone; Prof. Richard Lewis for $Na_V1.2$–1.5 and $Na_V1.8$ clones; Prof. Alan Goldin for the $Na_V1.6$ clone; and Prof. Frank Bosmans for the $Na_V1.7$ clone.

Declaration of Interests

The authors declare no competing interests.

References

1. Fry, B. G., Roelants, K., Champagne, D. E., Scheib, H., Tyndall, J. D., King, G. F., Nevalainen, T. J., Norman, J. A., Lewis, R. J., Norton, R. S., Renjifo, C. & de la Vega, R. C. (2009) The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms, *Annu Rev Genom Human Genet.* **10**, 483–511.

2. Jenner, R. & Undheim, E. (2017) *Venom: the secrets of nature's deadliest weapon*, Natural History Museum, London, UK.

3. King, G. F. (2015) *Venoms to drugs: venom as a source for the development of human therapeutics*, Royal Society of Chemistry, London, UK.

4. Casewell, N. R., Wuster, W., Vonk, F. J., Harrison, R. A. & Fry, B. G. (2013) Complex cocktails: the evolutionary novelty of venoms, *Trends Ecol Evol.* **28**, 219–229.

5. Van Iten, H., Marques, A. C., Leme, J. d. M., Pacheco, M. L. A. F. & Simões, M. G. (2014) Origin and early diversification of the phylum Cnidaria Verrill: major developments in the analysis of the taxon's Proterozoic–Cambrian history, *Palaeontology.* **57**, 677–690.

6. Frazao, B., Vasconcelos, V. & Antunes, A. (2012) Sea anemone (Cnidaria, Anthozoa, Actiniaria) toxins: an overview, *Mar Drugs.* **10**, 1812–1851.

7. Jouiaei, M., Yanagihara, A. A., Madio, B., Nevalainen, T. J., Alewood, P. F. & Fry, B. G. (2015) Ancient venom systems: a review on cnidaria toxins, *Toxins.* **7**, 2251–2271.

8. Mikov, A. N. & Kozlov, S. A. (2015) Structural features of cysteine-rich polypeptides from sea anemone venoms, *Russian J Bioorg Chem.* **41**, 455–466.

9. Logashina, Y. A., Solstad, R. G., Mineev, K. S., Korolkova, Y. V., Mosharova, I. V., Dyachenko, I. A., Palikov, V. A., Palikova, Y. A., Murashev, A. N., Arseniev, A. S., Kozlov, S. A., Stensvag, K., Haug, T. & Andreev, Y. A. (2017) New disulfide-stabilized fold provides sea anemone peptide to exhibit both antimicrobial and TRPA1 potentiating properties, *Toxins.* **9**, 154.

10. Madio, B., Undheim, E. A. B. & King, G. F. (2017) Revisiting venom of the sea anemone *Stichodactyla haddoni*: omics techniques reveal the complete toxin arsenal of a well-studied sea anemone genus, *J Proteomics.* **166**, 83–92.

11. Tarcha, E. J., Olsen, C. M., Probst, P., Peckham, D., Munoz-Elias, E. J., Kruger, J. G. & Iadonato, S. P. (2017) Safety and pharmacodynamics of dalazatide, a Kv1.3 channel inhibitor, in the treatment of plaque psoriasis: a randomized phase 1b trial, *PLoS One.* **12**, e0180762.

12. Orts, D. J. B., Moran, Y., Cologna, C. T., Peigneur, S., Madio, B., Praher, D., Quinton, L., De Pauw, E., Bicudo, J. E. P. W., Tytgat, J. & de Freitas, J. C. (2013) BcsTx3 is a founder of a novel sea anemone toxin family of potassium channel blocker, *FEBS J.* **280**, 4839–4852.

13. Malpezzi, E. L., de Freitas, J. C., Muramoto, K. & Kamiya, H. (1993) Characterization of peptides in sea anemone venom collected by a novel procedure, *Toxicon.* **31**, 853–864.

14. Fukuyama, Y., Iwamoto, S. & Tanaka, K. (2006) Rapid sequencing and disulfide mapping of peptides containing disulfide bonds by using 1,5-diaminonaphthalene as a reductive matrix, *J Mass Spectrom.* **41**, 191–201.

15. Undheim, E. A. B., Sunagar, K., Hamilton, B. R., Jones, A., Venter, D. J., Fry, B. G. & King, G. F. (2014) Multifunctional warheads: diversification of the toxin arsenal of centipedes via novel multidomain transcripts, *J Proteomics.* **102**, 1–10.

16. Mitchell, M. L., Hamilton, B. R., Madio, B., Morales, R. A. V., Tonkin-Hill, G. Q., Papenfuss, A. T., Purcell, A. W., King, G. F., Undheim, E. A. B. & Norton, R. S. (2017) The use of imaging mass spectrometry to study peptide toxin distribution in Australian sea anemones, *Austr J Chem.* **70**, 1235–1237.

17. Hale, J. E., Butler, J. P., Gelfanova, V., You, J. S. & Knierman, M. D. (2004) A simplified procedure for the reduction and alkylation of cysteine residues in proteins prior to proteolytic digestion and mass spectral analysis, *Anal Biochem.* **333**, 174-81.

18. Cock, P. J., Gruning, B. A., Paszkiewicz, K. & Pritchard, L. (2013) Galaxy tools and workflows for sequence analysis with applications in molecular plant pathology, *PeerJ.* **1**, e167.

19. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. & Madden, T. L. (2009) BLAST+: architecture and applications, *BMC Bioinformatics.* **10**, 421.

20. Katoh, K. & Standley, D. M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability, *Mol Biol Evol.* **30**, 772–780.

21. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies, *Mol Biol Evol.* **32**, 268–274.

22. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates, *Nat Meth.* **14**, 587–589.

23. Minh, B. Q., Nguyen, M. A. T. & von Haeseler, A. (2013) Ultrafast approximation for phylogenetic bootstrap, *Mol Biol Evol.* **30**, 1188–1195.

24. Cardoso, F. C., Dekan, Z., Smith, J. J., Deuis, J. R., Vetter, I., Herzig, V., Alewood, P. F., King, G. F. & Lewis, R. J. (2017) Modulatory features of the novel spider toxin mu-TRTX-Df1a isolated from the venom of the spider *Davus fasciatus*, *Br J Pharmacol.* **174**, 2528-2544.

25. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J. & Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline, *Proteins.* **59**, 687–696.

26. Shen, Y. & Bax, A. (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks, *J Biomol NMR.* **56**, 227–241.

27. Guntert, P. (2004) Automated NMR structure calculation with CYANA, *Methods Mol Biol.* **278**, 353–378.

28. Tytgat, J., Debont, T., Carmeliet, E. & Daenens, P. (1995) The alpha-dendrotoxin footprint on a mammalian potassium channel, *J Biol Chem.* **270**, 24776-81.

29. Edwards, I. A., Elliott, A. G., Kavanagh, A. M., Zuegg, J., Blaskovich, M. A. & Cooper, M. A. (2016) Contribution of amphipathicity and hydrophobicity to the antimicrobial activity and cytotoxicity of β-hairpin peptides, *ACS Infect Dis.* **2**, 442–450.

30. Huang, Y. H., Colgrave, M. L., Clark, R. J., Kotze, A. C. & Craik, D. J. (2010) Lysine-scanning mutagenesis reveals an amendable face of the cyclotide kalata B1 for the optimization of nematocidal activity, *J Biol Chem.* **285**, 10797–10805.

31. Torcato, I. M., Huang, Y. H., Franquelim, H. G., Gaspar, D., Craik, D. J., Castanho, M. A. & Troeira Henriques, S. (2013) Design and characterization of novel antimicrobial peptides, R-BP100 and RW-BP100, with activity against Gram-negative and Gram-positive bacteria, *Biochim Biophys Acta.* **1828**, 944–955.

32. Henriques, S. T., Huang, Y. H., Castanho, M. A., Bagatolli, L. A., Sonza, S., Tachedjian, G., Daly, N. L. & Craik, D. J. (2012) Phosphatidylethanolamine binding is a conserved feature of cyclotide-membrane interactions, *J Biol Chem.* **287**, 33629-43.

33. Henriques, S. T., Huang, Y. H., Rosengren, K. J., Franquelim, H. G., Carvalho, F. A., Johnson, A., Sonza, S., Tachedjian, G., Castanho, M. A., Daly, N. L. & Craik, D. J. (2011) Decoding the membrane activity of the cyclotide kalata B1: the importance of phosphatidylethanolamine phospholipids and lipid organization on hemolytic and anti-HIV activities, *J Biol Chem.* **286**, 24231-41.

34. Henriques, S. T., Pattenden, L. K., Aguilar, M. I. & Castanho, M. A. (2008) PrP(106-126) does not interact with membranes under physiological conditions, *Biophys J.* **95**, 1877-89.

35. Moran, Y., Genikhovich, G., Gordon, D., Wienkoop, S., Zenkert, C., Ozbek, S., Technau, U. & Gurevitz, M. (2012) Neurotoxin localization to ectodermal gland cells uncovers an alternative mechanism of venom delivery in sea anemones, *Proc Biol Sci.* **279**, 1351–1358.

36. Oliveira, J. S., Zaharenko, A. J., Ferreira, W. A., Jr., Konno, K., Shida, C. S., Richardson, M., Lucio, A. D., Beirao, P. S. & de Freitas, J. C. (2006) BcIV, a new paralyzing peptide obtained from the venom of the sea anemone *Bunodosoma caissarum*. A comparison with the Na+ channel toxin BcIII, *Biochim Biophys Acta.* **1764**, 1592-600.

37. Rodriguez, E., Barbeitos, M. S., Brugler, M. R., Crowley, L. M., Grajales, A., Gusmao, L., Haussermann, V., Reft, A. & Daly, M. (2014) Hidden among sea anemones: the first comprehensive phylogenetic reconstruction of the order Actiniaria (Cnidaria, Anthozoa, Hexacorallia) reveals a novel group of hexacorals, *PLoS One.* **9**, e96998.

38. Moran, Y., Weinberger, H., Sullivan, J. C., Reitzel, A. M., Finnerty, J. R. & Gurevitz, M. (2008) Concerted evolution of sea anemone neurotoxin genes is revealed through analysis of the *Nematostella vectensis* genome, *Mol Biol Evol.* **25**, 737–747.

39. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., 3rd, Snoeyink, J., Richardson, J. S. & Richardson, D. C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids, *Nucleic Acids Res.* **35**, 375–383.

40. Panteleev, P. V., Balandin, S. V., Ivanov, V. T. & Ovchinnikova, T. V. (2017) A therapeutic potential of animal beta-hairpin antimicrobial peptides, *Curr Med Chem.* **24**, 1724–1746.

41. Chagot, B., Pimentel, C., Dai, L., Pil, J., Tytgat, J., Nakajima, T., Corzo, G., Darbon, H. & Ferrat, G. (2005) An unusual fold for potassium channel blockers: NMR structure of three toxins from the scorpion *Opisthacanthus madagascariensis*, *Biochem J.* **388**, 263–271.

42. Osmakov, D. I., Kozlov, S. A., Andreev, Y. A., Koshelev, S. G., Sanamyan, N. P., Sanamyan, K. E., Dyachenko, I. A., Bondarenko, D. A., Murashev, A. N., Mineev, K. S., Arseniev, A. S. & Grishin, E. V. (2013) Sea anemone peptide with uncommon β-hairpin structure inhibits acid-sensing ion channel 3 (ASIC3) and reveals analgesic activity, *J Biol Chem.* **288**, 23116–23127.

43. Smith, J. J., Hill, J. M., Little, M. J., Nicholson, G. M., King, G. F. & Alewood, P. F. (2011) Unique scorpion toxin with a putative ancestral fold provides insight into evolution of the inhibitor cystine knot motif, *Proc Natl Acad Sci USA.* **108**, 10478–10483.

44. Yang, S., Yang, F., Wei, N., Hong, J., Li, B., Luo, L., Rong, M., Yarov-Yarovoy, V., Zheng, J., Wang, K. & Lai, R. (2015) A pain-inducing centipede toxin targets the heat activation machinery of nociceptor TRPV1, *Nat Commun.* **6**, 8297.

45. Henriques, S. T., Lawrence, N., Chaousis, S., Ravipati, A. S., Cheneval, O., Benfield, A. H., Elliott, A. G., Kavanagh, A. M., Cooper, M. A., Chan, L. Y., Huang, Y. H. & Craik, D. J. (2017) Redesigned spider peptide with improved antimicrobial and anticancer properties, *ACS Chem Biol.* **12**, 2324–2334.

588 46. Basulto, A., Perez, V. M., Noa, Y., Varela, C., Otero, A. J. & Pico, M. C. (2006)
589 Immunohistochemical targeting of sea anemone cytolysins on tentacles, mesenteric filaments and
590 isolated nematocysts of *Stichodactyla helianthus*, *J Exp Zool Part A, Comp Exp Biol.* **305**, 253–258.

591 47. Beckmann, A. & Ozbek, S. (2012) The nematocyst: a molecular map of the cnidarian stinging
592 organelle, *Int J Dev Biol.* **56**, 577–582.

593 48. Macrander, J., Brugler, M. R. & Daly, M. (2015) A RNA-seq approach to identify putative toxins
594 from acrorhagi in aggressive and non-aggressive *Anthopleura elegantissima* polyps, *BMC Genomics.*
595 **16**, 221.

596 49. Undheim, E. A. B., Hamilton, B. R., Kurniawan, N. D., Bowlay, G., Cribb, B. W., Merritt, D. J., Fry,
597 B. G., King, G. F. & Venter, D. J. (2015) Production and packaging of a biological arsenal: evolution of
598 centipede venoms under morphological constraint, *Proc Natl Acad Sci USA.* **112**, 4026–4031.

599 50. Moran, Y., Genikhovich, G., Gordon, D., Wienkoop, S., Zenkert, C., Ozbek, S., Technau, U. &
600 Gurevitz, M. (2012) Neurotoxin localization to ectodermal gland cells uncovers an alternative
601 mechanism of venom delivery in sea anemones, *Proc Biol Sci.* **279**, 1351–1358.

602 51. Chintiroglou, C. & Koukouras, A. (1992) The feeding habits of three Mediterranean sea anemone
603 species, *Anemonia viridis* (Forskål), *Actinia equina* (Linnaeus) and *Cereus pedunculatus* (Pennant),
604 *Helgol Mar Res.* **46**, 53–68.

605 52. Kruger, L. M. & Griffiths, C. L. (1996) Sources of nutrition in intertidal sea anemones from the
606 south-western Cape, South Africa, *S Afr J Zool.* **31**, 110–119.

607 53. Ottaway, J. R. (1978) Population ecology of the intertidal anemone *Actinia tenebrosa*: pedal
608 locomotion and intraspecific aggression, *Austr J Mar Freshwater Res.* **29**, 787–802.

609 54. Robson, S. A. & King, G. F. (2006) Domain architecture and structure of the bacterial cell division
610 protein DivIB, *Proc Natl Acad Sci USA.* **103**, 6700–6705.

611 55. Lummis, S. C., Beene, D. L., Lee, L. W., Lester, H. A., Broadhurst, R. W. & Dougherty, D. A. (2005)
612 *Cis-trans* isomerization at a proline opens the pore of a neurotransmitter-gated ion channel, *Nature.*
613 **438**, 248–252.

614 56. Andreotti, A. H. (2006) Opening the pore hinges on proline, *Nat Chem Biol.* **2**, 13–14.

615 57. Hille, B. (2001) *Ion channels of excitable membranes*, 3rd edn, Sinauer, Sunderland, MA, USA.

616 58. Bende, N. S., Dziemborowicz, S., Mobli, M., Herzig, V., Gilchrist, J., Wagner, J., Nicholson, G. M.,
617 King, G. F. & Bosmans, F. (2014) A distinct sodium channel voltage-sensor locus determines insect
618 selectivity of the spider toxin Dc1a, *Nat Commun.* **5**, 4350.

619 59. Sunagar, K. & Moran, Y. (2015) The rise and fall of an evolutionary innovation: Contrasting
620 strategies of venom evolution in ancient and young animals, *PLoS Genet.* **11**, e1005596.

621 60. Undheim, E. A. B., Mobli, M. & King, G. F. (2016) Toxin structures as evolutionary tools: using
622 conserved 3D folds to study the evolution of rapidly evolving peptides, *BioEssays.* **38**, 539–548.

623 61. Brown, D. D., Wensink, P. C. & Jordan, E. (1972) A comparison of the ribosomal DNAs of *Xenopus*
624 *laevis* and *Xenopus mulleri*: the evolution of tandem genes, *J Mol Biol.* **63**, 57–73.

625 62. Schüler, A. & Bornberg-Bauer, E. (2016) Evolution of protein domain repeats in metazoa, *Mol*
626 *Biol Evol.* **33**, 3170–3182.

627 63. Margres, M. J., Bigelow, A. T., Lemmon, E. M., Lemmon, A. R. & Rokyta, D. R. (2017) Selection to
628 increase expression, not sequence diversity, precedes gene family origin and expansion in
629 rattlesnake venom, *Genetics.* **206**, 1569–1580.

630 64. Honma, T., Hasegawa, Y., Ishida, M., Nagai, H., Nagashima, Y. & Shiomi, K. (2005) Isolation and
631 molecular cloning of novel peptide toxins from the sea anemone *Antheopsis maculata*, *Toxicon.* **45**,
632 33–41.

633

Figures and legends

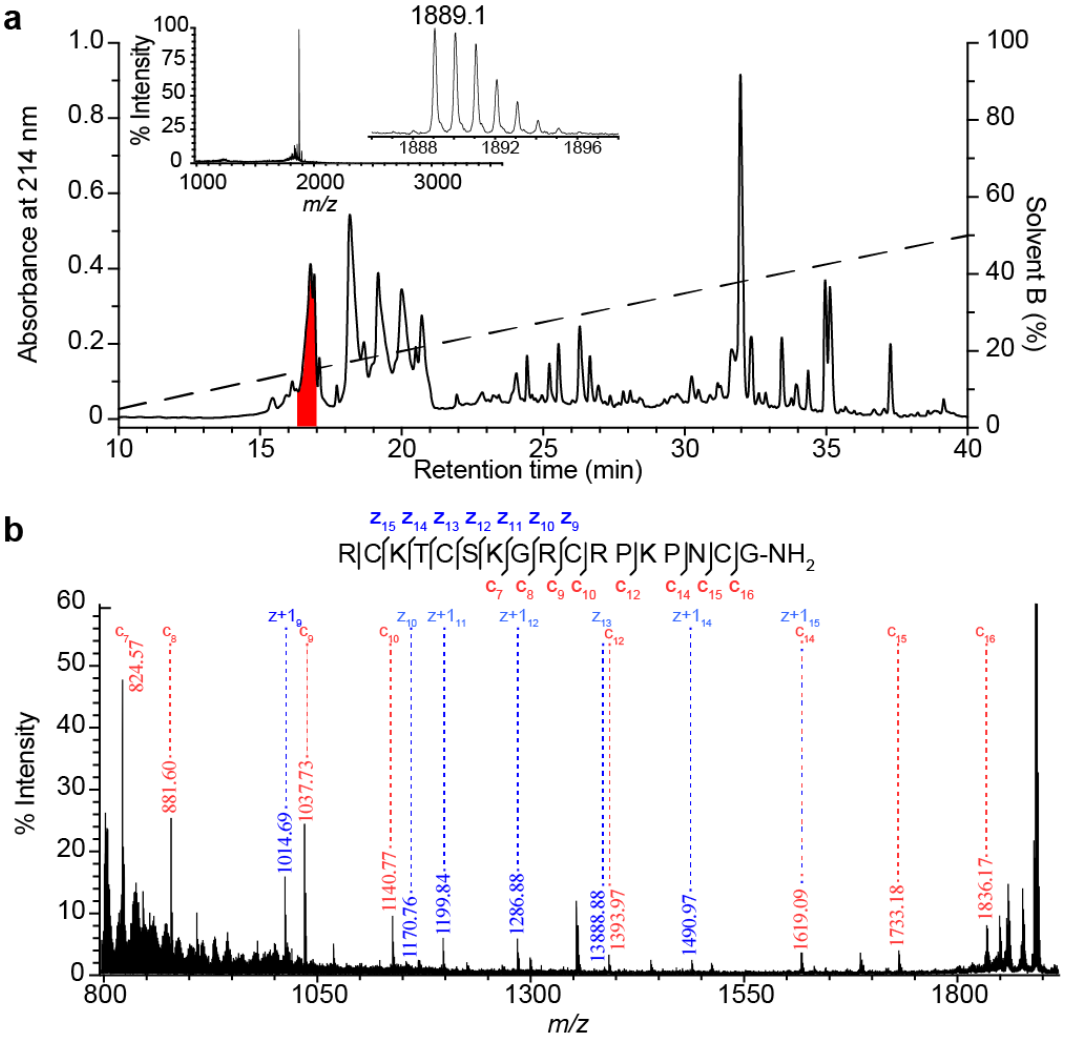


Fig. 1 Isolation and sequencing of Ate1a. (**a**) C_{18} RP-HPLC chromatogram showing fractionation of crude *A. tenebrosa* venom. The early-eluting peak containing Ate1a is highlighted in red. Inset shows average mass and isotope family for Ate1a. (**b**) *De novo* sequencing of Ate1a using ISD-MALDI MS

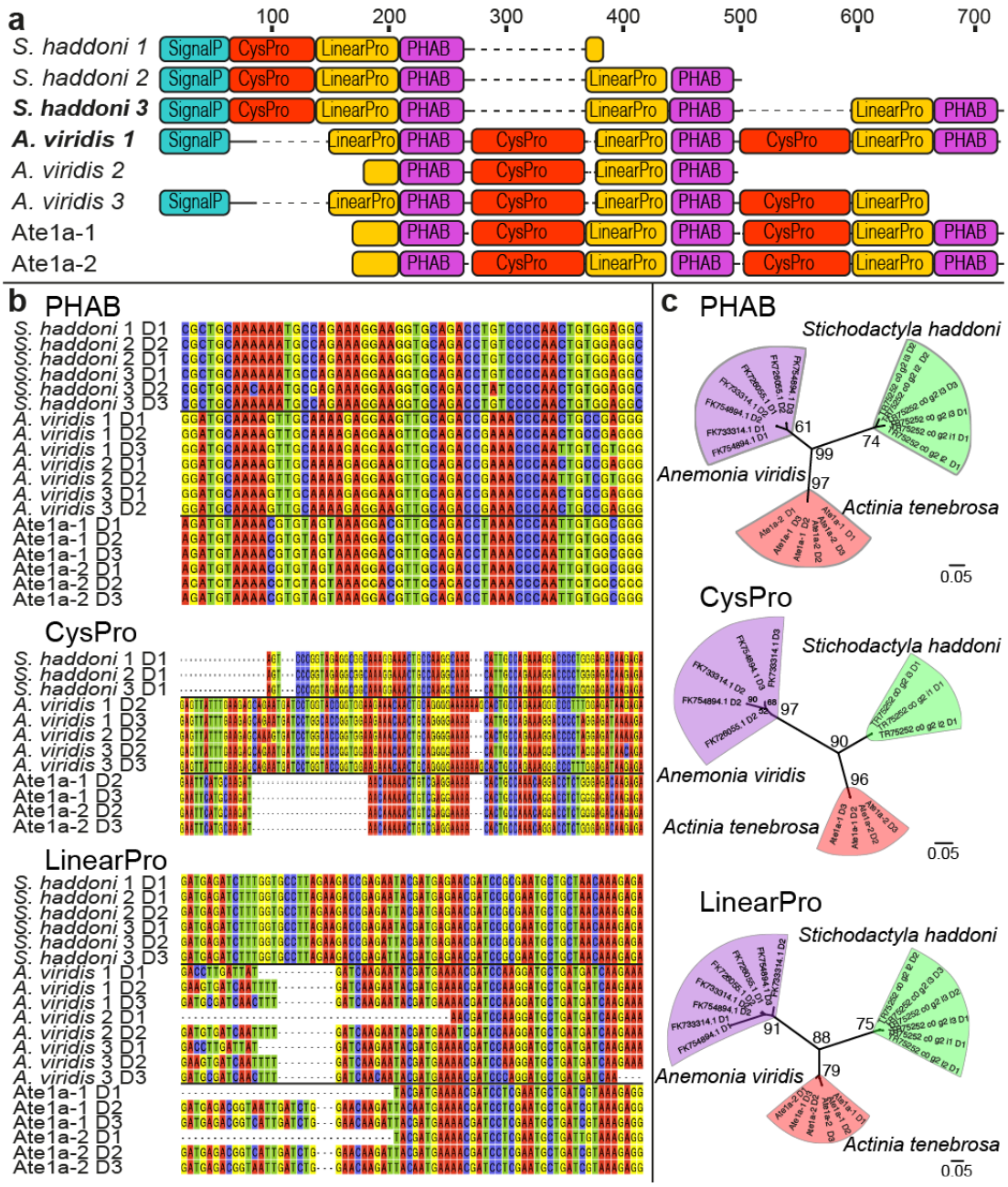


Fig. 2 Domain architecture and evolution of Ate1a precursors. (**a**) Domain architecture of Ate1a and Ate1a-like contigs. Prepropeptides are composed of a signal peptide (SignalP), one or two cysteine-containing propeptide domains (CysProP), and three cysteine-free propeptide domains (LinearProP) that each precede an Ate1a-like PHAB domain. (**b**) Nucleotide sequence alignments for each domain. (**c**) Maximum likelihood phylogenetic reconstructions for each domain. Bootstrap support values are shown at the nodes, while horizontal bars indicate genetic distance. Sequence accessions are for *S. haddoni* (1–3) TR75252_c0_g2_i1–3 and *A. viridis* (1) FK754894, (2) FK726055, and (3) FK733314. See also Figure S1

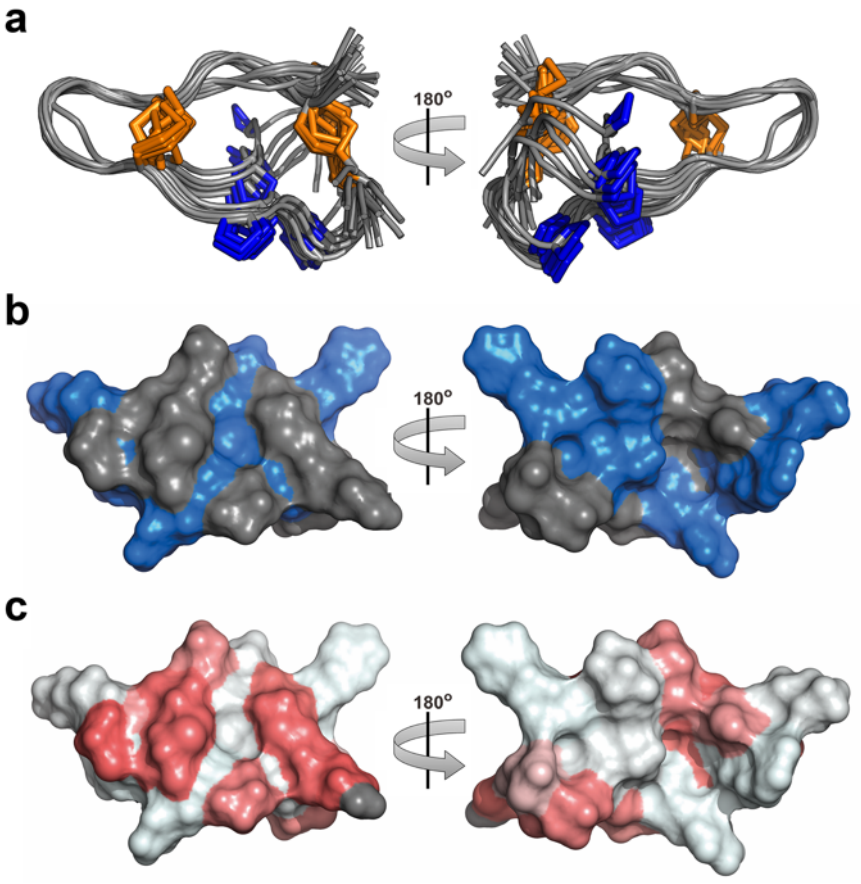


Fig. 3 3D structure of Ate1a. (**a**) Solution structure of Ate1a (ensemble of 20 structures; PDB code 6AZA). Disulfide bonds are highlighted in orange and proline side chains are shown in blue. (**b**) Surface representation of Ate1a with cationic and uncharged residues shown in blue and grey, respectively. (**c**) Surface representation of Ate1a showing relative hydrophobicity, which increases from white to red. See also Figure S2 and Table S1

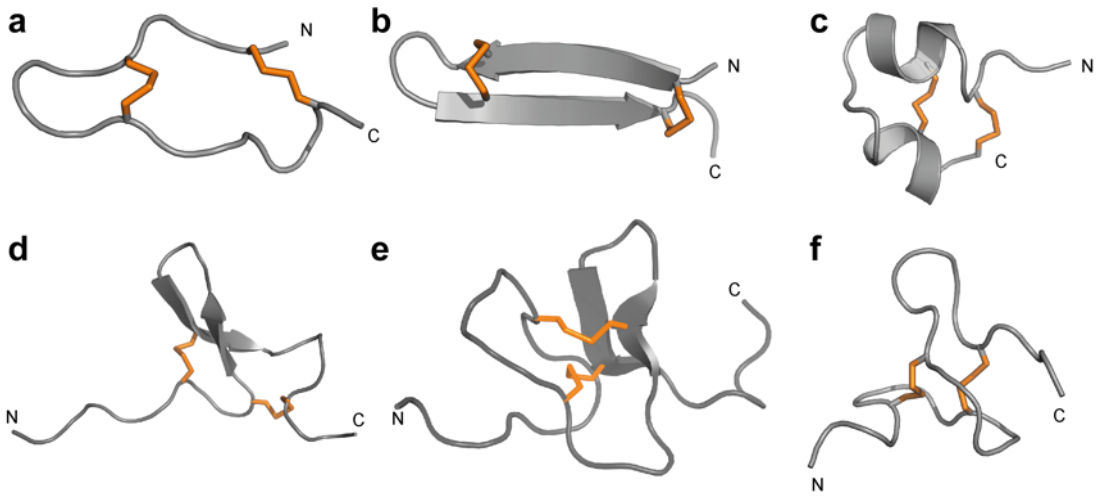
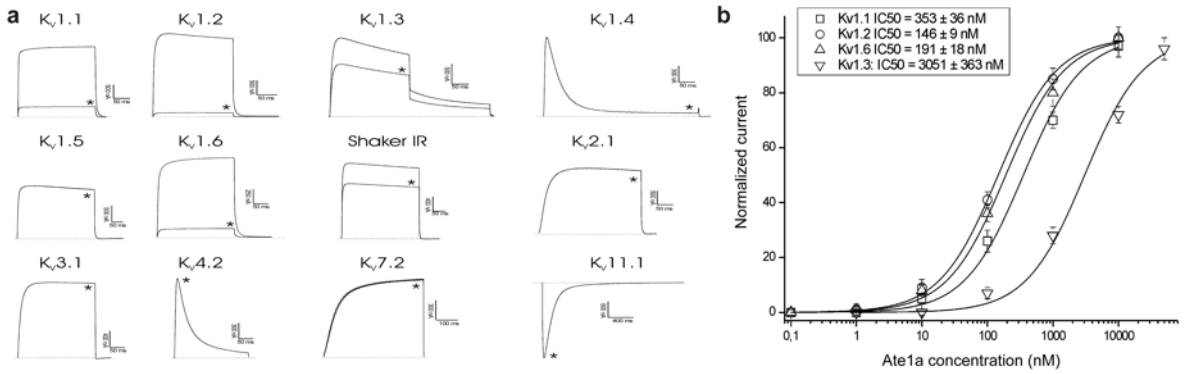


Fig. 4 Ate1a is the first member of the new PHAB fold. Comparison of the PHAB fold with other peptide folds containing two disulfide bonds and a similar number of residues (16-29 residues). Disulfide bonds are shown as orange tubes and N- and C-termini are labelled. (**a**) Ate1a; (**b**) β-hairpin fold represented by the spider peptide gomesin (PDB 1KFP); (**c**) CS α/α motif represented by the scorpion toxin κ-hefutoxin1 (PDB 1HP9); (**d**) Boundless β-hairpin motif represented by sea anemone toxin **π**-AnmTX Ugr 9a-1 (PDB: 2LZO); (**e**) Disulfide-directed hairpin represented by scorpion toxin U_1-Liotoxin-Lw1a (PDB 2KYJ); (**f**) Unstructured two-disulfide peptide fold represented by centipede toxin RhTx (PDB 2MVA)



670

671 **Fig. 5 Electrophysiological characterization of K_V isoforms inhibited by Ate1a.** (**a**) Representative
672 whole-cell current traces obtained from K_V channels expressed in *Xenopus* oocytes in the absence
673 (control) and presence (*) of 3 μM Ate1a. (**b**) Concentration-response curves obtained by plotting
674 current inhibition as a function of increasing Ate1a concentration. See also Figures S3–S5 and Table
675 S2

676

677

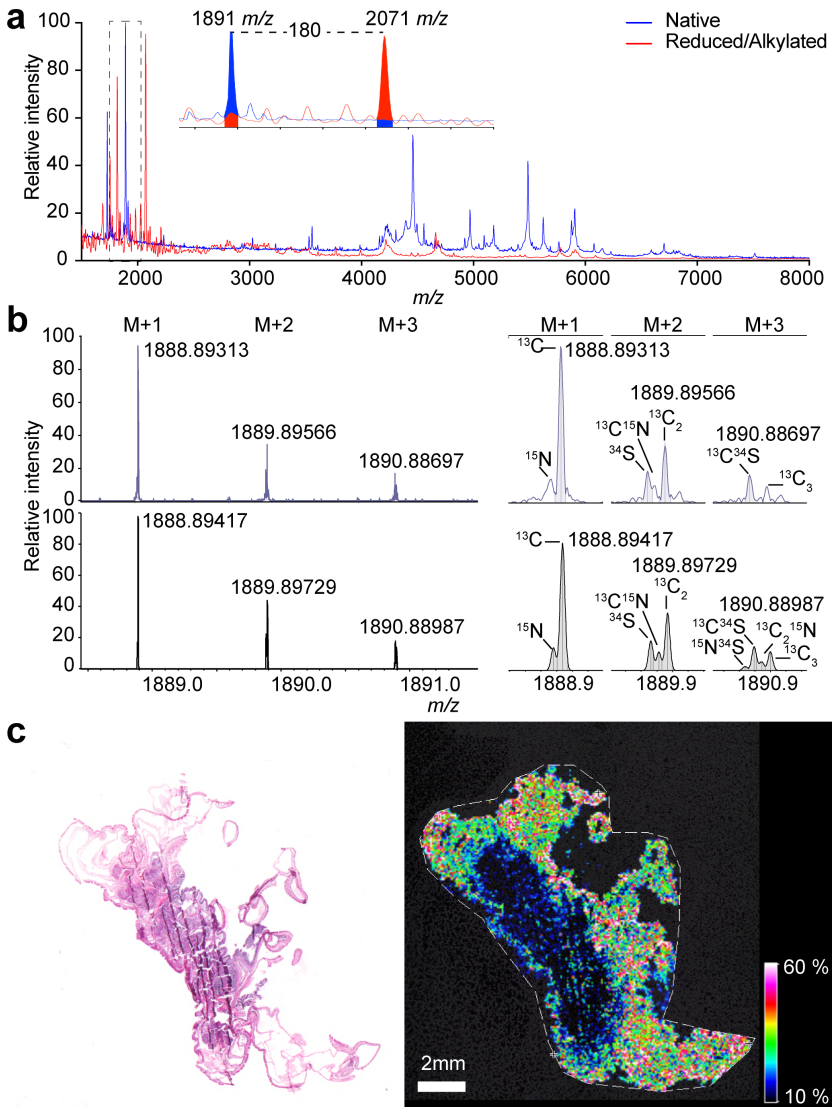


Fig. 6 Tissue distribution of Ate1a determined using MSI. (**a**) MSI linear positive mode spectra acquired from a cross-sectioned animal, with peaks corresponding to Ate1a filled in. The spectrum of native tissue is blue while the spectrum obtained after on-tissue gas-phase reduction and alkylation is shown in red. Inset shows a mass difference of 180 Da, corresponding to ethanolylation of four cysteine residues. (**b**) Ultra-high mass resolution analysis of the peak corresponding to Ate1a acquired by MALDI-FT-ICR-MSI at a resolution of 16,000,000 and resolving power at 1890 *m/z* of > 500,000, showing observed (top) and calculated (bottom) spectra. (**c**) Left: Histological image of the sea anemone section used for MSI experiments, stained with hematoxylin and eosin. Right: distribution of the peak corresponding to the average mass of Ate1a as observed by MALDI-TOF MSI. See also Supplemental videos S1 and S2